



University of Groningen

Toetsing van communicatieve vaardigheden

Vermeulen, Willem

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1993

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Vermeulen, W. (1993). Toetsing van communicatieve vaardigheden: constructie en evaluatie van gedragstoetsen voor professionele gespreksvaardigheden. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

In dit proefschrift worden de constructie en de evaluatie van gedragstoetsen voor professionele gespreksvaardigheden beschreven. In het eerste inleidende hoofdstuk worden de voorgeschiedenis en het ontstaan van het onderzoek behandeld. Het doel van het onderzoek was een beoordelingsinstrument te ontwikkelen voor intakegesprekken, zoals die bij de Gemeentelijke Sociale Dienst behoren te worden gevoerd.

In hoofdstuk 2 introduceren we de term "beroepsvaardigheden" om daarmee het complexe karakter van gespreksvaardigheden aan te geven en tevens de relatie met de beroepspraktijk te benadrukken. Aangevoerd wordt dat alleen met een integrale vaardigheidstoetsing de onderwijsdoelstellingen van de beroepsvaardigheid, en daarmee het achterliggende begrip, kunnen worden gedekt. Daarnaast wordt aangegeven dat er in het onderwijs verschillende ontwikkelingen gaande zijn die deze benadering ondersteunen.

Nadat we geconstateerd hebben dat het toetsen van gespreksvaardigheden in het MDGO (Middelbaar Dienstverlenings- en Gezondheidszorgonderwijs) nauwelijks gestandaardiseerd plaatsvindt, vinden we in het examenprogramma van de opleiding MDGO-SA/SD (MDGO, afdeling Sociale Arbeid, studierichting Sociale Dienstverlening) voldoende rechtvaardiging het intakegesprek als uitgangspunt voor ons onderzoek te nemen. Als beoordelingsmethode bij de beroepsvaardigheid "Het voeren van een intakegesprek" kiezen we voor de gedragstoets, omdat in deze toets de beroepsvaardigheid in zijn totaliteit aan de orde komt. Een probleem bij de gedragstoets vormt de standaardisatie. In de standaardisatie maken we een onderscheid tussen conditionele en onconditionele standaardisatie. Beide vormen van standaardisatie zijn in het onderzoek toegepast. Ten aanzien van de betrouwbaarheid van het simulatiespel blijkt uit de literatuur, dat hoge indices mogelijk zijn voor zowel de interne consistentie als voor de beoordelaarsovereenstemming. De validiteit van het simulatiespel blijkt matig te zijn.

In hoofdstuk 3 beschrijven we de constructie van de gedragstoets. De vaardigheden in de gedragstoets zijn ontleend aan de doelstellingen van het intakegesprek en worden verdeeld in drie clusters: 1. Gespreksstructurende vaardigheden, 2. Luistervaardigheden en 3. Vakinhoudelijke vaardigheden. Deze vaardigheden hebben vrijwel alle uitsluitend betrekking op waarneembaar gedrag.

Achtereenvolgens worden de werkzaamheden van een drietal werkgroepen bij de inventarisatie van de verschillende clusters van vaardigheden verantwoord. Voor de samenstelling van de definitieve lijst met Gespreksstructurende vaardigheden en Luistervaardigheden is in twee werkgroepen gebruik gemaakt van de methode van de rationele reconstructie: aan de hand van intakegesprekken op videoband werd in elk van de vijf bijeenkomsten op iteratieve wijze een voorlopige lijst met vaardigheden verbeterd tot een definitieve lijst. In de derde werkgroep zijn de casussen volgens het zogenaamde schema van opgavenkenmerken, alsmede de lijst met de daarbij behorende Vakinhoudelijke vaardigheden ontwikkeld. Deze drie lijsten van vaardigheden zijn vervolgens omgezet tot items. Ten slotte worden in het hoofdstuk de verschillende onderdelen van de gedragstoets, namelijk het simulatiespel, de afnameprocedure en de beoordelingsprocedure verantwoord.

In hoofdstuk 4 wordt het vooronderzoek beschreven. Daarin zijn aspecten van de betrouwbaarheid en de begripsvaliditeit van de gedragstoets alsmede de hanteerbaarheid ervan onderzocht. Als proefpersonen dienden 30 derdejaars leerlingen MDGO-SA/SD die

allen één intakegesprek voerden. Deze gesprekken werden op videoband opgenomen en beoordeeld door drie groepen beoordelaars, namelijk zes beoordelaars die te werk gingen volgens de procedure in de gedragstoets, zes docenten en zes werkvelddeskundigen. Deze laatste twee beoordelingen dienden als criteriumbeoordelingen. Ten behoeve van de vraag naar de hanteerbaarheid werd een enquête aan de leerlingen en de beoordelaars voorgelegd.

De resultaten van het vooronderzoek wijzen op een voldoende betrouwbaarheid van de gedragstoets: de interne consistentie is hoog en de beoordelaarsovereenstemming in eerste instantie veelbelovend. Verder is steun gevonden voor de begripsvaliditeit van de gedragstoets. Ook de hanteerbaarheid van de gedragstoets leverde geen echte problemen op. Naar aanleiding van het vooronderzoek is de gedragstoets op onderdelen bijgesteld.

In hoofdstuk 5 wordt de opzet van het hoofdonderzoek verantwoord. In vergelijking met het vooronderzoek is het hoofdonderzoek groter en breder van opzet. Ten eerste werden in het hoofdonderzoek drie steekproeven geformeerd: 25 eerstejaars, 25 derdejaars leerlingen MDGO-SA/SD en 25 dienstverleners van GSD's. Ten tweede voerden alle proefpersonen drie intakegesprekken. Er waren wederom drie groepen beoordelaars: getrainde beoordelaars die volgens de procedure in de gedragstoets werkten, docenten en deskundigen uit het werkveld. De laatste twee groepen leverden criteriumbeoordelingen; aan deze beoordelingen werd een derde criteriumbeoordeling toegevoegd, namelijk het oordeel door de cliënt. Ten slotte zijn van de leerlingen MDGO-SA/SD enkele rapport- en examencijfers opgevraagd. De onderzoeksvragen richtten zich weer op aspecten van de betrouwbaarheid en validiteit van de toets met het accent op het laatste aspect.

In hoofdstuk 6 bespreken we de resultaten van het hoofdonderzoek. Daarbij blijkt dat de betrouwbaarheid van de gedragstoets goed is: de interne consistentie is hoog, terwijl ook de beoordelaarsovereenstemming aanzienlijk is. Ten aanzien van de inhoudsvaliditeit wordt geconstateerd dat de gedragstoets voldoende representatief is voor de betreffende beroepsvaardigheid. Begripsvaliditeit van de gedragstoets blijkt ten eerste uit de hoge correlaties tussen de gedragstoets en de drie criteriumbeoordelingen: het docentoordeel, het werkveldoordeel en het oordeel van de cliënt. De begripsvaliditeit wordt verder ondersteund omdat de gedragstoets blijkt te kunnen differentiëren tussen de drie groepen proefpersonen waarbij een verschil in beheersingsgraad mag worden verondersteld. Een ondersteuning voor de conclusie dat met de gedragstoets vaardigheden worden gemeten, die thans in de betreffende opleiding nog niet worden getoetst, zien we in het ontbreken van een correlationeel verband tussen de gedragstoets en de belangrijkste examenonderdelen van die opleiding.

Enkele generaliseerbaarheidsstudies wijzen het volgende uit. Ten eerste wordt aangetoond dat bij gedragstoetsing het inzetten van meer casussen meer winst in betrouwbaarheid oplevert dan wanneer meer beoordelaars worden ingeschakeld. Ten tweede wijst één van de studies uit dat de casussen niet als onderling inwisselbaar kunnen worden beschouwd en ten derde is vastgesteld dat beoordelaars die de procedure volgen die bij de gedragstoets hoort wel als onderling inwisselbaar kunnen worden gezien. Ten slotte is gebleken dat de indeling van de drie clusters van vaardigheden van de gedragstoets empirische ondersteuning vindt.

In hoofdstuk 7 hebben we de gedragstoets op een aantal punten vergeleken met de beoordeling van intakegesprekken in de praktijk. Afgemeten aan het werkveldoordeel is de beoordelingsprocedure bij de gedragstoets even valide als het docentoordeel. De beoordelingsprocedure bij de gedragstoets is evenwel betrouwbaarder dan het docentoordeel. Illustratief voor het verschil in betrouwbaarheid is de verhouding tussen het aantal beoordelaars dat voor elk van de respectievelijke beoordelingswijzen nodig is om een beoordelaarsovereenstemming van bijvoorbeeld 0.80 (ICC) te verwachten. Deze verhouding tussen de beoordelingsprocedure bij de gedragstoets, het docentoordeel en het werkveldoordeel bedraagt respectievelijk ongeveer $1 : 3\frac{1}{2} : 2$.

Een vergelijking op het gebied van de efficiëntie, in termen van hanteerbaarheid, kosten

en tijd blijkt niet mogelijk. Het is niet mogelijk om procedures voor gesprekken te ontwikkelen voor de jaarlijkse productie van de toets, de tijd die nodig is om een intakegesprek op te nemen, en de tijd die nodig is om op ongeveer anderhalf uur te komen.

In hoofdstuk 8 wordt de betrouwbaarheid van de gedragstoets op grond van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen in gesprekken met de cliënt, vergeleken met de betrouwbaarheid van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen op het werkveldoordeel. Het aantal beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen op het werkveldoordeel is aanzienlijk lager dan het aantal beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen op het intakegesprek.

In hoofdstuk 9 wordt de betrouwbaarheid van de gedragstoets op grond van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen in gesprekken met de cliënt, vergeleken met de betrouwbaarheid van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen op het werkveldoordeel. Het aantal beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen op het werkveldoordeel is aanzienlijk lager dan het aantal beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen op het intakegesprek.

Wat de betrouwbaarheid van de gedragstoets op grond van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen in gesprekken met de cliënt, betreft, wordt geconstateerd dat de betrouwbaarheid van de gedragstoets op grond van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen in gesprekken met de cliënt, voldoende is voor de betreffende beroepsvaardigheid.

Verder wordt geconstateerd dat de betrouwbaarheid van de gedragstoets op grond van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen in gesprekken met de cliënt, voldoende is voor de betreffende beroepsvaardigheid.

Na een vergelijking van de betrouwbaarheid van de gedragstoets op grond van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen in gesprekken met de cliënt, met de betrouwbaarheid van de gedragstoets op grond van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen op het werkveldoordeel, wordt geconstateerd dat de betrouwbaarheid van de gedragstoets op grond van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen in gesprekken met de cliënt, voldoende is voor de betreffende beroepsvaardigheid.

Een resultaat met betrekking tot de betrouwbaarheid van de gedragstoets op grond van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen in gesprekken met de cliënt, is dat de betrouwbaarheid van de gedragstoets op grond van de beoordelingen van de docenten, de leerlingen en de werkvelddeskundigen in gesprekken met de cliënt, voldoende is voor de betreffende beroepsvaardigheid.

en tijd blijkt niet mogelijk wegens het ontbreken van gestandaardiseerde beoordelingsprocedures voor gespreksvoering in het MDGO. Geschat wordt dat de ontwikkelingskosten voor de jaarlijkse productie van tien gedragstoetsen op ongeveer f 30.000,- komen. Voor de tijd die nodig is om een gedragstoets bij één kandidaat af te nemen, inclusief het opnemen van het intakegesprek op videoband en het beoordelen van het gesprek op deze videoband moet op ongeveer anderhalf uur worden gerekend.

In hoofdstuk 8 wordt ingegaan op het gebruik van de gedragstoets in de schoolpraktijk. Benadrukt wordt dat de toets primair is ontwikkeld voor resultaatbepaling. Daarnaast biedt de gedragstoets op grond van de gedetailleerde beoordeling mogelijkheden om in de klas het onderwijs in gespreksvoering sturing te geven. Waar de duurzaamheid van de gedragstoets redelijk positief blijkt uit te vallen, levert de geheimhouding van de casussen een probleem op. Ten slotte wordt een advies voor normering gegeven. Het blijkt dat dit advies, dat gebaseerd is op het werkveldoordeel en op de betreffende onderwijsdoelstellingen tot aanzienlijke aantallen onvoldoendes leidt.

In hoofdstuk 9 wordt aangegeven dat de hoofddoelstelling van het onderzoek, namelijk het construeren van een instrument waarmee intakegesprekken op een zo objectief mogelijke manier kunnen worden beoordeeld, bereikt is. Aan de hand van de drie eisen die aan beoordelingsinstrumenten worden gesteld, namelijk betrouwbaarheid, validiteit en efficiëntie, wordt bovenstaande conclusie onderbouwd door de volgende resultaten.

Wat de betrouwbaarheid van de gedragstoets betreft, is vastgesteld dat met de beoordeling door één beoordelaar een acceptabele betrouwbaarheid realiseerbaar is. De begripsvaliditeit van de gedragstoets wordt vooral aangetoond door de grote mate van overeenstemming tussen de beoordelingsprocedure bij de gedragstoets, het oordeel van docenten en deskundigen uit het werkveld over de criteria voor het voeren van een professioneel intakegesprek.

Verder wordt geconstateerd dat de gedragstoets door de docenten en de leerlingen goed hanteerbaar wordt geacht, maar dat een meer omvattende evaluatie van dit onderwerp op termijn nodig is. Het gebruik van de gedragstoets blijkt duur te zijn en kost veel tijd.

Na een vergelijking tussen de beoordelingsprocedure bij de gedragstoets en de bestaande beoordelingswijze wordt vastgesteld dat, wanneer de beoordeling door één docent geschiedt, de eerste beoordelingswijze zowel betrouwbaarder is als meer begripsvaliditeit bezit. De conclusie wordt dan ook getrokken dat de gedragstoets geschikt is om als examentoets voor het voeren van intakegesprekken te worden gebruikt.

Een resultaat met een meer algemeen karakter is het feit dat de in dit onderzoek ontwikkelde methoden en procedures ook bruikbaar zijn voor de constructie van beoordelingsinstrumenten voor andere beroepenvelden waarin gespreksvoering een belangrijke plaats inneemt. Het hoofdstuk eindigt met het maken van enkele kanttekeningen bij het onderzoek en met het geven van enkele suggesties voor vervolgonderzoek.